

0590
1224

#10

ENTERED



OIPE

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/039,183A

DATE: 01/07/2003

TIME: 16:13:47

Input Set : A:\06132.037002.SEQLIST.TXT

Output Set: N:\CRF4\01072003\J039183A.raw

```

4 <110> APPLICANT: Kleanthous, Harold
5     Lissolo, Ling
6     Tomb, Jean-Francois
7     Miller, Charles
8     Al-Garawi, Amal
10 <120> TITLE OF INVENTION: Helicobacter GHPO 1360 and GHPO 750
11     Polypeptides and Corresponding Polynucleotide Molecules
14 <130> FILE REFERENCE: 06132/037002
16 <140> CURRENT APPLICATION NUMBER: US 10/039,183A
C--> 17 <141> CURRENT FILING DATE: 2002-12-13
19 <150> PRIOR APPLICATION NUMBER: US 08/831,310
20 <151> PRIOR FILING DATE: 1997-04-01
22 <160> NUMBER OF SEQ ID NOS: 18
24 <170> SOFTWARE: FastSEQ for Windows Version 4.0
26 <210> SEQ ID NO: 1
27 <211> LENGTH: 1149
28 <212> TYPE: DNA
29 <213> ORGANISM: Helicobacter pylori
31 <220> FEATURE:
32 <221> NAME/KEY: CDS
33 <222> LOCATION: (106)...(1002)
35 <221> NAME/KEY: sig_peptide
36 <222> LOCATION: (106)...(166)
W--> 38 <400> 1
39 ttactcttta atgtgagttt tctgtgtcat gatagctgat tttgttttaa atttgctata 60
40 atgtgaattt aatgatgaaa attagtttag agtggagaac acaca atg aaa aaa aat 117
41                                     Met Lys Lys Asn
42                                     -20
44 atc tta aat tta gcg tta gtg ggt gcg ttg agc acg tcg ttt ttg atg 165
45 Ile Leu Asn Leu Ala Leu Val Gly Ala Leu Ser Thr Ser Phe Leu Met
46 -15 -10 -5
48 gct aag ccg gct cat aac gca aat aac gct acg cat aac acg aaa aaa 213
49 Ala Lys Pro Ala His Asn Ala Asn Asn Ala Thr His Asn Thr Lys Lys
50 1 5 10 15
52 acg act gat tct tca gca ggc gtg tta gcg aca gtg gat ggc aga cct 261
53 Thr Thr Asp Ser Ala Gly Val Leu Ala Thr Val Asp Gly Arg Pro
54 20 25 30
56 atc act aaa agc gat ttt gac atg att aag caa cga aat cct aat ttt 309
57 Ile Thr Lys Ser Asp Phe Asp Met Ile Lys Gln Arg Asn Pro Asn Phe
58 35 40 45
60 gat ttt gac aag ctt aaa gag aaa gaa aaa gaa gcc ttg att gat caa 357
61 Asp Phe Asp Lys Leu Lys Glu Lys Glu Lys Glu Ala Leu Ile Asp Gln
62 50 55 60

```

RAW SEQUENCE LISTING

DATE: 01/07/2003

PATENT APPLICATION: US/10/039,183A

TIME: 16:13:47

Input Set : A:\06132.037002.SEQLIST.TXT

Output Set: N:\CRF4\01072003\J039183A.raw

```

64 gct att cgc acc gcc ctt gta gaa aat gaa gct aaa acc gag aaa ttg 405
65 Ala Ile Arg Thr Ala Leu Val Glu Asn Glu Ala Lys Thr Glu Lys Leu
66 65 70 75 80
68 gac agc act cca gaa ttt aaa gcg atg atg gaa gcg gtt aaa aaa cag 453
69 Asp Ser Thr Pro Glu Phe Lys Ala Met Met Glu Ala Val Lys Lys Gln
70 85 90 95
72 gct tta gtg gaa ttt tgg gct aaa aaa cag gct gaa gaa gtg aaa aaa 501
73 Ala Leu Val Glu Phe Trp Ala Lys Lys Gln Ala Glu Glu Val Lys Lys
74 100 105 110
76 gtc caa atc cca gaa aaa gaa atg caa gat ttt tac aac gct aac aaa 549
77 Val Gln Ile Pro Glu Lys Glu Met Gln Asp Phe Tyr Asn Ala Asn Lys
78 115 120 125
80 gat cag ctt ttt gtc aag caa gaa gcc cat gct agg cat att tta gtg 597
81 Asp Gln Leu Phe Val Lys Gln Glu Ala His Ala Arg His Ile Leu Val
82 130 135 140
84 aaa acc gaa gat gag gct aaa cgg att att tct gag att gac aaa cag 645
85 Lys Thr Glu Asp Glu Ala Lys Arg Ile Ile Ser Glu Ile Asp Lys Gln
86 145 150 155 160
88 cca aag gct aaa aaa gaa gct aaa ttc att gag tta gcc aat cgg gat 693
89 Pro Lys Ala Lys Lys Glu Ala Lys Phe Ile Glu Leu Ala Asn Arg Asp
90 165 170 175
92 acg att gat cct aac agc aag aac gcg caa aat ggc ggt gat ttg ggg 741
93 Thr Ile Asp Pro Asn Ser Lys Asn Ala Gln Asn Gly Gly Asp Leu Gly
94 180 185 190
96 aaa ttc caa aag aac caa atg gct ccg gat ttt tct aaa gcc gct ttc 789
97 Lys Phe Gln Lys Asn Gln Met Ala Pro Asp Phe Ser Lys Ala Ala Phe
98 195 200 205
100 gct tta act cct ggg gat tac act aaa acc cct gtt aaa aca gag ttt 837
101 Ala Leu Thr Pro Gly Asp Tyr Thr Lys Thr Pro Val Lys Thr Glu Phe
102 210 215 220
104 ggt tat cat att atc tat ttg att tct aaa gat agc cct gta act tat 885
105 Gly Tyr His Ile Ile Tyr Leu Ile Ser Lys Asp Ser Pro Val Thr Tyr
106 225 230 235 240
108 act tat gaa cag gct aaa cct acc att aag ggg atg tta caa gaa aag 933
109 Thr Tyr Glu Gln Ala Lys Pro Thr Ile Lys Gly Met Leu Gln Glu Lys
110 245 250 255
112 ctt ttc caa gaa cgc atg aat caa cgc att gag gaa cta aga aag cac 981
113 Leu Phe Gln Glu Arg Met Asn Gln Arg Ile Glu Glu Leu Arg Lys His
114 260 265 270
116 gct aaa att gtt atc aac aag taattgatga ggtgttatca tgtagttaa 1032
117 Ala Lys Ile Val Ile Asn Lys
118 275
120 aggcaatgaa attttattga aagcccataa agaaggttat ggggtggggg cgtttaattt 1092
121 cgtgaatttt gaaatgctaa acgctatttt tgaagcagga aatgaggaaa attcccc 1149
123 <210> SEQ ID NO: 2
124 <211> LENGTH: 299
125 <212> TYPE: PRT
126 <213> ORGANISM: Helicobacter pylori
128 <220> FEATURE:

```

RAW SEQUENCE LISTING

DATE: 01/07/2003

PATENT APPLICATION: US/10/039,183A

TIME: 16:13:47

Input Set : A:\06132.037002.SEQLIST.TXT

Output Set: N:\CRF4\01072003\J039183A.raw

```

129 <221> NAME/KEY: SIGNAL
130 <222> LOCATION: (1)...(20)
132 <400> SEQUENCE: 2
133 Met Lys Lys Asn Ile Leu Asn Leu Ala Leu Val Gly Ala Leu Ser Thr
134 -20 -15 -10 -5
135 Ser Phe Leu Met Ala Lys Pro Ala His Asn Ala Asn Asn Ala Thr His
136 1 5 10
137 Asn Thr Lys Lys Thr Thr Asp Ser Ser Ala Gly Val Leu Ala Thr Val
138 15 20 25
139 Asp Gly Arg Pro Ile Thr Lys Ser Asp Phe Asp Met Ile Lys Gln Arg
140 30 35 40
141 Asn Pro Asn Phe Asp Phe Asp Lys Leu Lys Glu Lys Glu Lys Glu Ala
142 45 50 55 60
143 Leu Ile Asp Gln Ala Ile Arg Thr Ala Leu Val Glu Asn Glu Ala Lys
144 65 70 75
145 Thr Glu Lys Leu Asp Ser Thr Pro Glu Phe Lys Ala Met Met Glu Ala
146 80 85 90
147 Val Lys Lys Gln Ala Leu Val Glu Phe Trp Ala Lys Lys Gln Ala Glu
148 95 100 105
149 Glu Val Lys Lys Val Gln Ile Pro Glu Lys Glu Met Gln Asp Phe Tyr
150 110 115 120
151 Asn Ala Asn Lys Asp Gln Leu Phe Val Lys Gln Glu Ala His Ala Arg
152 125 130 135 140
153 His Ile Leu Val Lys Thr Glu Asp Glu Ala Lys Arg Ile Ile Ser Glu
154 145 150 155
155 Ile Asp Lys Gln Pro Lys Ala Lys Lys Glu Ala Lys Phe Ile Glu Leu
156 160 165 170
157 Ala Asn Arg Asp Thr Ile Asp Pro Asn Ser Lys Asn Ala Gln Asn Gly
158 175 180 185
159 Gly Asp Leu Gly Lys Phe Gln Lys Asn Gln Met Ala Pro Asp Phe Ser
160 190 195 200
161 Lys Ala Ala Phe Ala Leu Thr Pro Gly Asp Tyr Thr Lys Thr Pro Val
162 205 210 215 220
163 Lys Thr Glu Phe Gly Tyr His Ile Ile Tyr Leu Ile Ser Lys Asp Ser
164 225 230 235
165 Pro Val Thr Tyr Thr Tyr Glu Gln Ala Lys Pro Thr Ile Lys Gly Met
166 240 245 250
167 Leu Gln Glu Lys Leu Phe Gln Glu Arg Met Asn Gln Arg Ile Glu Glu
168 255 260 265
169 Leu Arg Lys His Ala Lys Ile Val Ile Asn Lys
170 270 275
173 <210> SEQ ID NO: 3
174 <211> LENGTH: 1448
175 <212> TYPE: DNA
176 <213> ORGANISM: Helicobacter pylori
178 <220> FEATURE:
179 <221> NAME/KEY: CDS
180 <222> LOCATION: (118)...(1314)
182 <400> SEQUENCE: 3

```

RAW SEQUENCE LISTING

DATE: 01/07/2003

PATENT APPLICATION: US/10/039,183A

TIME: 16:13:47

Input Set : A:\06132.037002.SEQLIST.TXT

Output Set: N:\CRF4\01072003\J039183A.raw

```

183 ctcttgaatg gcgataagac aaaaatgtct taaattttgt ggtagcattt aggaatactt 60
184 aggattttgt ttagtataat tctaaaatcc atttcaaaaa attaaggaga aatacaa atg 120
185                                         Met
186                                         1
188 gca aaa gaa aag ttt aac aga act aag ccg cat gtt aat att gga acc 168
189 Ala Lys Glu Lys Phe Asn Arg Thr Lys Pro His Val Asn Ile Gly Thr
190                                         5          10          15
192 att ggg cat gta gac cat ggt aaa acg act ttg agt gca gcg att tca 216
193 Ile Gly His Val Asp His Gly Lys Thr Thr Leu Ser Ala Ala Ile Ser
194                                         20          25          30
196 gcg gtg ctt tct ttg aaa ggt ctt gca gaa atg aaa gac tat gat aat 264
197 Ala Val Leu Ser Leu Lys Gly Leu Ala Glu Met Lys Asp Tyr Asp Asn
198                                         35          40          45
200 att gat aac gcc cct gaa gaa aaa gaa aga ggg atc act atc gct act 312
201 Ile Asp Asn Ala Pro Glu Glu Lys Glu Arg Gly Ile Thr Ile Ala Thr
202 50                                         55          60          65
204 tct cac att gaa tat gag act gaa aac aga cac tat gcg cat gtg gat 360
205 Ser His Ile Glu Tyr Glu Thr Glu Asn Arg His Tyr Ala His Val Asp
206                                         70          75          80
208 tgc cca gga cac gct gac tat gta aaa aac atg atc acc ggt gcg gcg 408
209 Cys Pro Gly His Ala Asp Tyr Val Lys Asn Met Ile Thr Gly Ala Ala
210                                         85          90          95
212 caa atg gac gga gcg att ttg gtt gtt tct gca gct gat ggc cct atg 456
213 Gln Met Asp Gly Ala Ile Leu Val Ser Ala Ala Asp Gly Pro Met
214                                         100         105         110
216 cct caa act agg gag cat atc tta ttg tct cgt caa gta ggc gtg cct 504
217 Pro Gln Thr Arg Glu His Ile Leu Leu Ser Arg Gln Val Gly Val Pro
218                                         115         120         125
220 cac atc gtt gtt ttc tta aac aaa caa gac atg gta gat gac caa gaa 552
221 His Ile Val Val Phe Leu Asn Lys Gln Asp Met Val Asp Asp Gln Glu
222 130                                         135         140         145
224 ttg tta gaa ctt gta gaa atg gaa gtg cgc gaa ttg ttg agc gcg tat 600
225 Leu Leu Glu Leu Val Glu Met Glu Val Arg Glu Leu Leu Ser Ala Tyr
226                                         150         155         160
228 gaa ttt cct ggc gat gac act cct atc gta gcg ggt tca gct tta aga 648
229 Glu Phe Pro Gly Asp Asp Thr Pro Ile Val Ala Gly Ser Ala Leu Arg
230                                         165         170         175
232 gct tta gaa gaa gca aag gct ggt aat gtg ggt gaa tgg ggt gaa aaa 696
233 Ala Leu Glu Glu Ala Lys Ala Gly Asn Val Gly Glu Trp Gly Glu Lys
234                                         180         185         190
236 gtg ctt aaa ctt atg gct gaa gtg gat gcc tat atc cct act cca gaa 744
237 Val Leu Lys Leu Met Ala Glu Val Asp Ala Tyr Ile Pro Thr Pro Glu
238                                         195         200         205
240 aga gac act gaa aaa act ttc ttg atg ccg gtt gaa gat gtg ttc tct 792
241 Arg Asp Thr Glu Lys Thr Phe Leu Met Pro Val Glu Asp Val Phe Ser
242 210                                         215         220         225
244 att gcg ggt aga ggg act gtg gtt aca ggt agg att gaa aga ggc gtg 840
245 Ile Ala Gly Arg Gly Thr Val Val Thr Gly Arg Ile Glu Arg Gly Val
246                                         230         235         240

```

RAW SEQUENCE LISTING

DATE: 01/07/2003

PATENT APPLICATION: US/10/039,183A

TIME: 16:13:47

Input Set : A:\06132.037002.SEQLIST.TXT

Output Set: N:\CRF4\01072003\J039183A.raw

```

248 gtg aaa gta ggc gat gaa gtg gaa atc gtt ggt atc aga cct aca caa      888
249 Val Lys Val Gly Asp Glu Val Glu Ile Val Gly Ile Arg Pro Thr Gln
250          245          250          255
252 aaa acg act gta acc ggt gta gaa atg ttt agg aaa gag ttg gaa aaa      936
253 Lys Thr Thr Val Thr Gly Val Glu Met Phe Arg Lys Glu Leu Glu Lys
254          260          265          270
256 ggt gaa gcc ggc gat aat gtg ggc gtg ctt ttg aga gga act aaa aaa      984
257 Gly Glu Ala Gly Asp Asn Val Gly Val Leu Leu Arg Gly Thr Lys Lys
258          275          280          285
260 gaa gaa gtg gaa cgc ggt atg gtt cta tgc aaa cca ggt tct atc act      1032
261 Glu Glu Val Glu Arg Gly Met Val Leu Cys Lys Pro Gly Ser Ile Thr
262 290          295          300          305
264 ccg cac aag aaa ttt gag gga gaa att tat gtc ctt tct aaa gaa gaa      1080
265 Pro His Lys Lys Phe Glu Gly Glu Ile Tyr Val Leu Ser Lys Glu Glu
266          310          315          320
268 ggc ggg aga cac act cca ttc ttc acc aat tac cgc ccg caa ttc tat      1128
269 Gly Gly Arg His Thr Pro Phe Phe Thr Asn Tyr Arg Pro Gln Phe Tyr
270          325          330          335
272 gtg cgc aca act gat gtg act ggc tct atc acc ctt cct gaa ggc gta      1176
273 Val Arg Thr Thr Asp Val Thr Gly Ser Ile Thr Leu Pro Glu Gly Val
274          340          345          350
276 gaa atg gtt atg cct ggc gat aat gtg aaa atc act gta gag ttg att      1224
277 Glu Met Val Met Pro Gly Asp Asn Val Lys Ile Thr Val Glu Leu Ile
278          355          360          365
280 agc cct gtt gcg tta gag ttg gga act aaa ttt gcg att cgt gaa ggc      1272
281 Ser Pro Val Ala Leu Glu Leu Gly Thr Lys Phe Ala Ile Arg Glu Gly
282 370          375          380          385
284 ggt agg acc gtt ggt gct ggt gtt gtg agc aat att att gaa      1314
285 Gly Arg Thr Val Gly Ala Gly Val Val Ser Asn Ile Ile Glu
286          390          395
288 taatattagc aaaaagagag ttaccataaa gggtcattat gaaagttaaa atagggttga 1374
289 agtgtttctga ttgtgaagat atcaattaca gcacaaccaa gaacgctaaa actaacctg 1434
290 aaaaactgga gctt      1448
292 <210> SEQ ID NO: 4
293 <211> LENGTH: 399
294 <212> TYPE: PRT
295 <213> ORGANISM: Helicobacter pylori
297 <400> SEQUENCE: 4
298 Met Ala Lys Glu Lys Phe Asn Arg Thr Lys Pro His Val Asn Ile Gly
299 1          5          10          15
300 Thr Ile Gly His Val Asp His Gly Lys Thr Thr Leu Ser Ala Ala Ile
301          20          25          30
302 Ser Ala Val Leu Ser Leu Lys Gly Leu Ala Glu Met Lys Asp Tyr Asp
303          35          40          45
304 Asn Ile Asp Asn Ala Pro Glu Glu Lys Glu Arg Gly Ile Thr Ile Ala
305          50          55          60
306 Thr Ser His Ile Glu Tyr Glu Thr Glu Asn Arg His Tyr Ala His Val
307 65          70          75          80
308 Asp Cys Pro Gly His Ala Asp Tyr Val Lys Asn Met Ile Thr Gly Ala

```

VERIFICATION SUMMARY

DATE: 01/07/2003

PATENT APPLICATION: US/10/039,183A

TIME: 16:13:48

Input Set : A:\06132.037002.SEQLIST.TXT

Output Set: N:\CRF4\01072003\J039183A.raw

L:17 M:271 C: Current Filing Date differs, Replaced Current Filing Date

L:38 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ ID#:1